5. SURVEY METHODOLOGY

The general methods used to develop and deploy the GSPS were based on previously tested methods used for the BRFSS. Detailed information on the BRFSS methodology is available on the BRFSS website at http://www.cdc.gov/BRFSS/.

5.1 SURVEY DESIGN

The GSPS utilized a complex survey design, which included stratification and clustering. GSPS was designed to sample the coastal populations of the four states directly exposed to the *Deepwater Horizon* oil spill (Louisiana, Mississippi, Alabama, and Florida). The 2010 coastal population consisted of 3.3 million persons, 1.9 million in Louisiana (57% of the coastal population), 0.4 million in Mississippi (11%), 0.4 million in Alabama (11%), and 0.7 million in Florida (21%). The largest cities included in the coastal sample were New Orleans, Louisiana and Mobile, Alabama.

Initially, people were included in the coastal population sample if they resided in one of 25 coastal counties that lie within 32 miles of areas closed to fishing following the oil spill. Beginning in May 2011, CDC added counties in the four States that were further from the gulf to the survey sample, in order to allow comparison of results from the Gulf Coast counties to noncoastal counties. The noncoastal counties include all counties within the four States that were not part of the 25 coastal county areas. Also beginning in May 2011, CDC added a Spanish version of the questionnaire (Spanish was chosen as the second language of administration because Census population data indicate that the coastal areas have higher percentages of Hispanic/Latino populations than any other non-English-speaking residents) and also added cell phone respondents to the sample. Inclusion of cell phone in the frame was deemed especially important for purposes of coverage and validity because of the rising proportion of households in the United States that no longer use a landline phone.

In all, 38,361 interviews were conducted (32,813 on landlines and 5,548 on cell phones); 15,028 from Louisiana's coastal parishes, 3,209 from Mississippi coastal counties, 4,533 from Alabama, and 5,177 from Florida and 10,414 in noncoastal areas in all four states. A total of 122 surveys were conducted in Spanish; all others were conducted in English. Dataset users may identify language of interview by noting the variable LANG1 in the dataset.

The following table summarizes the number of completed interviews by phone type and residence location.

Table 1: Number of Completed Surveys by Telephone Type and Location: Gulf States Population Survey (GSPS), December 2010 – December 2011

Type of Telephone	Total number of surveys completed	Number completed in coastal counties	Number completed in noncoastal counties
Landline	32,813	24,468	8,345
Cell Phone	5,548	3,479	2,069
Total	38,361	27,947	10,414

5.2 DISPROPORTIONATE STRATIFIED SAMPLE (DSS)

Like the BRFSS, the GSPS used a disproportionate stratified sample (DSS) random-digit-dialing methodology for landline phones. DSS design increases efficiency of random-digit dialing phone surveys by dividing listed phone numbers into those expected to belong mostly to households (high-density) and those expected to contain a smaller proportion of household numbers (medium-density). Available phone numbers in each county were divided into blocks of 100 (e.g., 404-555-1100 through 404-555-1199 would comprise a block). Within each block high density was defined as all numbers containing a known attribution to a household. Medium density included all remaining numbers in the block (i.e., those not in the published household listing). High and medium density numbers were sampled at a 1.5:1 ratio (high to medium) to obtain a probability sample of all households with landline telephones in all parts of the coastal area. Cell phone numbers were not subject to DSS because data on their location is less specific and less reliable than that associated with landline phone numbers.

5.3 DATA WEIGHTING

Data weighting for the GSPS used standard procedures in place for the BRFSS. Following data collection, the GSPS data were cleaned of errors and subjected to weighting procedures. Landline and cell phone datasets were merged, and data were weighted to adjusted U.S. census population estimates by county, using age, race, sex, geographic location, and phone type (landline or cell phone). Data were weighted using county level characteristics (resulting in the variable _GSPSWT).

Data weighting is conducted to reduce bias in the sample and: 1) correct for differences in the probability of selection due to nonresponse and non-coverage; 2) adjust for differences in sex, age, race/ethnicity between the sample and the entire population, and; 3) permit generalization of survey findings to the population.

In brief, GSPS used a two-step weighting process, as in BRFSS: 1) design weighting to account for variations within households, and; 2) raking (also known as iterative

proportional fitting) to match known characteristics of the population to those in the sample. Design factors included the number of residential telephones in a household, the number of adults in a household, and geographic or density stratification. The GSPS used raking procedures for weighting to population totals. The following variables were used in the raking weighting process:

- Age
- Sex
- Race
- Marital Status
- Locational strata
- Phone source/ownership

Table 2 describes the data weighting variables included in the GSPS dataset.

Table 2: Data Weighting Variables: Gulf States Population Survey (GSPS), December 2010 – December 2011

Variable	Description
_GEOSTR	The geographic stratification variable. This identifies the location of the interview for weighting purposes.
_STSTR	The strata variable which should be used to account for strata when analyzing data in complex sampling statistical procedures.
_DENSTR2	The variable indicating density stratum for landline respondents.
_GSPSWT	Final adjusted weight assigned to each respondent at the county level. In some cases counties were combined before weighting due to lower number of respondents. These counties were: Cameron and Jefferson in Louisiana, Plaquemines and St. Bernard in Louisiana, Assumption and St. Mary in Louisiana and Hancock and Harrison in Mississippi.
_PSU	The PSU is used as a cluster and is unique within each of the states. It is a proxy for the phone number. Since landline phone numbers may have more than one adult, the household phone is the cluster. _PSU should be used as the cluster variable in complex survey analyses of the data.